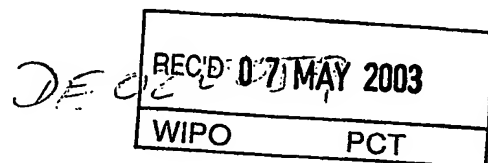


**PRIORITY  
DOCUMENT**

SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)



**Prioritätsbescheinigung über die Einreichung  
einer Patentanmeldung**

**Aktenzeichen:** 102 09 928.6

**Anmeldetag:** 07. März 2002

**Anmelder/Inhaber:** Philips Corporate Intellectual Property GmbH,  
Hamburg/DE

**Bezeichnung:** Verfahren zum Betrieb eines Sprach-Dialogsystems

**IPC:** G 10 L, G 06 F

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 25. Februar 2003  
Deutsches Patent- und Markenamt  
Der Präsident  
Im Auftrag



## ZUSAMMENFASSUNG

### Verfahren zum Betrieb eines Sprach-Dialogsystems

- Es wird ein Verfahren zum Betrieb eines Sprach-Dialogsystems (1) beschrieben. Zur Steuerung des Dialogs zum Auswählen eines von mehreren Diensten (9, 10) des Dialogsystems (1) durch den Nutzer wird eine Datenbank (6) verwendet, mit einer hierarchischen Datenstruktur (DS) mit einer Mehrzahl von Knotenpunkten (K) und einer Mehrzahl von Pfaden (P) zur Verbindung der Knotenpunkte (K) untereinander und zur Verbindung von Knotenpunkten (K) mit Dienstobjekten (D), welche jeweils an einem Ende eines Pfades (P) angeordnet sind. Die Dienstobjekte (D) repräsentieren die zur Verfügung stehenden Dienste (9, 10). Die Knotenpunkte (K) repräsentieren jeweils Kategorien, in welche darunter angeordnete weitere Kategorien und/oder die Dienste klassifiziert sind. Zumindest zu einem Teil der Dienstobjekte (D) und/oder Knotenpunkte (K) führt eine Mehrzahl von Pfaden (P). Jedem Knotenpunkt (K) und jedem Dienstobjekt (D) ist außerdem mindestens ein Schlüsselwort (S) zugeordnet. Aus einer Spracheingabe des Nutzers werden Suchworte extrahiert und auf Basis der Suchworte wird eine Anzahl von Kandidaten-Knotenpunkten (K) und/oder Kandidaten-Dienstobjekten (D) gesucht, deren zugeordnete Schlüsselworte (S) gemäß einem vorgegebenen Akzeptanzkriterium mit den Suchworten übereinstimmen. Das Suchverfahren wird in mehreren Suchschritten durchgeführt, bis die Anzahl der ermittelten Kandidaten-Knotenpunkte (K) und/oder Kandidaten-Dienstobjekte (D) oberhalb einer vorgegebenen Minimalanzahl und unterhalb einer vorgegebenen Maximalanzahl liegt. Es wird dann ein Sprachausgabe-Menü generiert und die durch die ermittelten Kandidaten-Knotenpunkte (K) und/oder Kandidaten-Dienstobjekte (D) repräsentierten Kategorien und/oder Dienste (9, 10) werden angesagt. Darüber hinaus wird ein entsprechendes Sprach-Dialogsystem (1) beschrieben.

BESCHREIBUNG

## Verfahren zum Betrieb eines Sprach-Dialogsystems

- Die Erfindung betrifft ein Verfahren zum Betrieb eines Sprach-Dialogsystems, welches mit einem Nutzer unter Verwendung einer Spracherkennungseinrichtung und einer
- 5 Sprachausgabereinrichtung kommuniziert, wobei dem Nutzer verschiedene Dienste innerhalb des Sprach-Dialogsystems oder durch Vermittlung des Sprach-Dialogsystems zur Verfügung stehen und vom Nutzer innerhalb eines mit dem Sprach-Dialogsystem geführten Dialogs auswählbar sind. Dabei wird zur Steuerung des Dialogs zum
- 10 Auswählen eines Dienstes durch den Nutzer eine Datenbank mit einer hierarchischen Datenstruktur verwendet, mit einer Mehrzahl von Knotenpunkten und einer Mehrzahl von Pfaden zur Verbindung der Knotenpunkte untereinander und zur Verbindung von Knotenpunkten mit Dienstobjekten, welche jeweils an einem Ende eines Pfads in der Datenstruktur angeordnet sind. Die Dienstobjekte repräsentieren dabei die zur
- 15 Verfügung stehenden Dienste, und die Knotenpunkte repräsentieren jeweils Kategorien, in welche wiederum andere Kategorien und/oder Dienste, die durch die innerhalb der hierarchischen Datenstruktur auf einer Ebene unter dem jeweiligen Knotenpunkt angeordneten weiteren Knotenpunkte oder Dienstobjekte repräsentiert werden, klassifiziert sind. Darüber hinaus betrifft die Erfindung ein entsprechendes automatisches Sprach-Dialogsystem und ein Computerprogramm mit Programmcodemitteln, um das
- 20 Verfahren auszuführen.

- Sprach-Dialogsysteme, die mit einem Nutzer unter Verwendung einer Spracherkennungs- und einer Sprachausgabereinrichtung kommunizieren, sind bereits seit längerem bekannt. Es handelt sich hierbei um sogenannte sprachgesteuerte, automatische
- 25 Systeme, die oft auch als Sprachapplikationen bezeichnet werden. Sofern es sich bei dem Sprach-Dialogsystem für den Nutzer um ein Mittel handelt, mit dem dieser den Zugang zu verschiedenen Diensten erlangt, handelt es sich um ein sogenanntes Voice Portal. Das Sprach-Dialogsystem kann spezielle Terminals aufweisen, an denen sich der Nutzer

befinden muss, um mit dem Sprach-Dialogsystem kommunizieren zu können, wie beispielsweise ein stationäres Auskunftssystem in einem Flughafen oder dergleichen. Häufig weisen derartige Sprach-Dialogsysteme jedoch einen Anschluss an ein öffentliches Kommunikationsnetz auf, so dass die Sprach-Dialogsysteme zum Beispiel

5 mittels eines normalen Telefons, eines Mobilfunkgeräts oder eines PCs mit Telefoniefunktion etc. genutzt werden können. Ein Beispiel für diese Sprach-Dialogsysteme sind automatische Anrufbeantwortungs- und Auskunftssysteme, wie sie inzwischen beispielsweise von einigen größeren Firmen, Organisationen und Ämtern verwendet werden, um einen Anrufer möglichst schnell und komfortabel mit der gewünschten

10 Information zu versorgen oder mit einer Stelle zu verbinden, welche für die speziellen Wünsche des Anrufers zuständig ist. Weitere Beispiele hierfür sind die automatische Telefonauskunft, wie sie bereits von einigen Telefongesellschaften eingesetzt wird, eine automatische Fahrplan- bzw. Flugplanauskunft oder ein Informationsdienst mit allgemeinen Veranstaltungshinweisen, beispielsweise mit Kino- und Theaterprogrammen,

15 für eine bestimmte Region. Einige der Sprach-Dialogsysteme bieten über das reine Angebot, Informationen für den Nutzer bereit zu halten oder zu suchen und bei Bedarf an den Nutzer zu übermitteln, auch zusätzliche Dienste wie beispielsweise einen Reservierungsdienst für Sitzplätze im Zug oder Flugzeug oder Hotelzimmer, einen Bezahlservice oder einen Warenbestellservice.

20 Hierbei kann der Nutzer auch – beispielsweise mittels einer Dialogvermittlung (auch „Call-Transfer“ genannt) an einen externen, d.h. nicht zum Sprach-Dialogsystem gehörenden – Dienst oder eine Person vermittelt werden. Der Begriff „Dienst“ umfasst aber im Sinne dieser Schrift ausdrücklich nicht nur einen komplexen Dienst wie eine Auskunftssystem, eine Vermittlungseinrichtung, einen Reservierungsdienst etc., sondern

25 es kann sich hierbei auch lediglich um eine einzelne Information handeln, die innerhalb des Sprach-Dialogsystems an den Nutzer als eine „Dienstleistung“ ausgegeben wird, beispielsweise die Ausgabe einer erfragten Telefonnummer oder das Abspielen eines Bandes mit Veranstaltungshinweisen. Im Prinzip können, ähnlich wie beispielsweise im Internet, dem Nutzer folglich beliebige Dienste über ein solches Sprach-Dialogsystem

30 angeboten werden. Bei einem Sprach-Dialogsystem besteht dabei der Vorteil, dass der

Nutzer lediglich ein normales Telefon bzw. ein Mobilfunkgerät benötigt, um die Dienste in Anspruch zu nehmen.

- Zur Auswahl eines bestimmten Dienstes des Sprach-Dialogsystems durch den Nutzer wird in der Praxis derzeit üblicherweise nach dem eingangs genannten Verfahren vorgegangen, bei dem die einzelnen Dienste in einer festgelegten, hierarchischen Datenstruktur angeordnet sind, welche wie ein Entscheidungsbaum aufgebaut ist. Der Dialog zwischen dem Nutzer und dem Sprach-Dialogsystem beginnt dabei an einem Startpunkt an der Spitze der „Baumstruktur“ und verläuft entlang eines Pfades bzw. Zweigs über mehrere Knotenpunkte, welche jeweils eine bestimmte Kategorie von Diensten repräsentieren, bis zum Ende eines Pfades, an dem sich dann ein „Dienstobjekt“ befindet, welches den entsprechenden Dienst repräsentiert. Unter dem Begriff „Dienstobjekt“ ist dabei im Sinne dieser Schrift ein beliebiges Datenobjekt, ein Softwaremodul o. Ä. zu verstehen, welches den Dienst selber darstellt und/oder Informationen über den Dienst enthält. Hierbei kann es sich beispielsweise um Informationen, in welcher Form der Dienst aufzurufen ist, um eine Adresse des Dienstes bzw. des betreffenden Softwaremoduls oder um Informationen zur Durchführung eines Call-Transfers oder Ähnliches handeln.
- Die Knotenpunkte, welche jeweils die Kategorien repräsentieren, befinden sich auf verschiedenen Ebenen, wobei jeweils die Knotenpunkte auf einer höheren Ebene Kategorien repräsentieren, in die wiederum die Kategorien eingeordnet sind, die zu den in der darunter liegenden Ebene befindlichen Knotenpunkten gehören und somit bezüglich der darüber liegenden Kategorie sozusagen die Unterkategorien bilden.
- Ein anschauliches Beispiel hierfür ist ein Sprach-Dialogsystem, welches verschiedene Auskunftsdienste anbietet, u. a. beispielsweise einen Weiterbericht und Veranstaltungshinweise. Bei diesem Beispiel könnte von einem zentralen Knotenpunkt aus in die Kategorien „Wetter“ und „Veranstaltungshinweise“ unterteilt werden. Unterhalb dieser Kategorien befinden sich dann jeweils weitere Kategorien, beispielsweise unter der

Kategorie „Wetter“ eine Kategorie „aktuelles Reisewetter“ und eine Kategorie „Wettervorhersage“ und unter der Kategorie „Veranstaltungshinweise“ die Kategorien „Kino“, „Theater“ und „Vorträge“ etc. Unter diesen einzelnen Kategorien befinden sich dann wiederum weitere Kategorien, wie beispielsweise unter der Kategorie „Reisewetter“ die einzelnen Regionen, für die das Wetter abfragbar ist, oder unter der Kategorie „Theater“ die einzelnen Theater in einer Stadt. Der Benutzer kann dann innerhalb des Dialogs einen Dienst auswählen, indem er ausgehend vom Startpunkt zunächst die Kategorien der oberen Ebene angeboten bekommt und zur Auswahl einer Kategorie aufgefordert wird. Dies kann beispielsweise durch eine Sprachausgabe des Systems (im Folgenden auch „Prompt“ genannt) auf folgende Weise geschehen: „Wenn Sie sich für Veranstaltungshinweise interessieren, sagen Sie bitte 'Veranstaltungen', wenn Sie sich für den Wetterbericht interessieren, sagen Sie bitte 'Wetter'.“ Je nach Antwort des Benutzers wird dann ein erneuter Prompt vom Dialogsystem erzeugt, beispielsweise nach Auswahl des Wetterberichts die Ansage: „Wenn Sie das aktuelle Reisewetter wünschen, sagen Sie 'Reise', wenn Sie den Wetterbericht für die nächsten Tage haben möchten, sagen Sie 'Wetterbericht'“ usw.

Es ist klar, dass mit zunehmender Komplexität der Sprach-Dialogsysteme und zunehmender Anzahl der Dienste innerhalb des Sprach-Dialogsystems die Baumstruktur entsprechend immer komplizierter wird und immer mehr Ebenen und somit auch mehr Knotenpunkte zur Verzweigung der einzelnen Pfade hinzukommen müssen. Um zu einem bestimmten Dienst zu gelangen, muss der Benutzer dabei zunächst die gesamte Datenstruktur vom Startpunkt bis zum Ende des jeweiligen Pfades durchlaufen und eine Vielzahl von Abfragen des Dialogsystems beantworten. Dieses Vorgehen ist dementsprechend zeitaufwändig, langweilig und unkomfortabel für den Benutzer. Darüber hinaus ist eine einmal festgelegte, starre Kategorieaufteilung zwangsläufig nicht für jeden der verschiedenen Nutzer intuitiv gut bedenkbar, so dass es leicht zu Fehlentscheidungen des Nutzers an einem Knotenpunkt kommen kann. Der Nutzer wird dann zu einem falschen Dienst geleitet und muss unter Umständen noch einmal mit dem gesamten Dialog von vorne beginnen. Ist die Komplexität des gesamten Systems zu groß, führt

dies dazu, dass das Sprach-Dialogsystem wegen dieser Unbequemlichkeiten von den Nutzern nicht mehr in Anspruch genommen wird.

5 Es ist Aufgabe der vorliegenden Erfindung, ein verbessertes Verfahren zum Betrieb eines Sprach-Dialogsystems der eingangs genannten Art und ein entsprechendes Sprach-Dialogsystem zu schaffen, welches es ermöglicht, dass der Nutzer jederzeit auf einfache und schnelle Weise den gewünschten Dienst findet.

10 Diese Aufgabe wird durch ein Verfahren der eingangs genannten Art gelöst, bei dem zumindest zu einem Teil der Dienstobjekte und/oder Knotenpunkte innerhalb der Datenstruktur eine Mehrzahl von verschiedenen Pfaden führt und bei dem jedem Knotenpunkt und jedem Dienstobjekt ein oder mehrere Schlüsselworte zugeordnet sind. Erfindungsgemäß werden bei einem Empfang einer Spracheingabe des Nutzers aus dieser Spracheingabe Suchworte extrahiert und auf Basis der Suchworte eine Anzahl von 15 Kandidaten-Knotenpunkten und/oder Kandidaten-Dienstobjekten gesucht, deren zugeordnete Schlüsselworte gemäß einem vorgegebenen Akzeptanzkriterium mit den Suchworten übereinstimmen. Dabei wird eine Suche in mehreren Suchschritten durchgeführt, bis nach einem Suchschritt die Anzahl der ermittelten Kandidaten-Knotenpunkte und/oder Kandidaten-Dienstobjekte oberhalb einer vorgegebenen 20 Minimalzahl und unterhalb einer vorgegebenen Maximalzahl liegt. Es wird dann mittels der Sprachausgabeeinrichtung ein Sprachausgabe-Menü zur Ansage der durch die ermittelten Kandidaten-Knotenpunkte und/oder Kandidaten-Dienstobjekte repräsentierten Kategorien und/oder Dienste an den Nutzer zur Auswahl einer bestimmten Kategorie oder eines bestimmten Dienstes generiert.

25 Vorrichtungsmäßig wird die Aufgabe durch ein automatisches Sprach-Dialogsystem mit einer Spracherkennungseinrichtung und einer Sprachausgabeeinrichtung zur Kommunikation mit einem Nutzer sowie mit mehreren vom Nutzer ausführbaren Diensten innerhalb des Sprach-Dialogsystems und/oder mit Mitteln zur Vermittlung des Nutzers an 30 vom Nutzer auswählbare Dienste durch das Sprach-Dialogsystem gelöst, wobei dieses

Sprach-Dialogsystem eine Dialogkontrolleinrichtung zur Steuerung des Dialogs zum Auswählen eines Dienstes durch den Nutzer und eine Datenbank mit einer entsprechenden oben genannten hierarchischen Datenstruktur mit einer Mehrzahl von Knotenpunkten und einer Mehrzahl von Pfaden zur Verbindung der Knotenpunkte untereinander und zur Verbindung von Knotenpunkten mit Dienstobjekten aufweist, welche jeweils an einem Ende eines Pfades in der Datenstruktur angeordnet sind. Dabei repräsentieren die Dienstobjekte die zur Verfügung stehenden Dienste, und die Knotenpunkte repräsentieren jeweils die Kategorien, in welche wiederum andere Kategorien und/oder Dienste, die durch die innerhalb der hierarchischen Datenstruktur auf einer Ebene unter dem jeweiligen Knotenpunkt angeordneten Knotenpunkte oder Dienstobjekte repräsentiert werden, klassifiziert sind. Erfindungsgemäß führt dabei zumindest zu einem Teil der Dienstobjekte und/oder Knotenpunkte innerhalb der Datenstruktur eine Mehrzahl von verschiedenen Pfaden. Außerdem sind jedem Knotenpunkt und jedem Dienstobjekt der Datenbank ein oder mehrere Schlüsselworte zugeordnet. Des Weiteren muss das erfindungsgemäße Sprach-Dialogsystem eine Analyseeinheit, um bei einem Empfang einer Spracheingabe des Nutzers aus dieser Spracheingabe Suchworte zu extrahieren, und eine Sucheinheit aufweisen, um auf Basis der Suchworte eine Anzahl von Kandidaten-Knotenpunkten und/oder Kandidaten-Dienstobjekte innerhalb der Datenbank zu suchen, deren zugeordnete Schlüsselworte gemäß einem vorgegebenen Akzeptanzkriterium mit den Suchworten übereinstimmen, wobei die Sucheinheit derart aufgebaut ist, dass sie eine Suche in mehreren Suchschritten durchführt, bis nach einem Suchschritt die Anzahl der ermittelten Kandidaten-Knotenpunkte und/oder Kandidaten-Dienstobjekte oberhalb einer vorgegebenen Minimalanzahl und unterhalb einer vorgegebenen Maximalanzahl liegt. Schließlich braucht das erfindungsgemäße Sprach-Dialogsystem eine Promptgenerierungseinheit, um nach einem erfolgreichen Suchschritt ein Sprachausgabe-Menü als Prompt zu generieren, um so dem Nutzer die durch die ermittelten Kandidaten-Knotenpunkte und/oder Kandidaten-Dienstobjekte repräsentierten Kategorien und/oder Dienste zur Auswahl einer bestimmten Kategorie oder eines bestimmten Dienstes mittels der Sprachausgabeeinrichtung anzusagen.



- Durch die Zuordnung von Schlüsselworten zu jedem Knotenpunkt und jedem Dienstobjekt – bzw. zu den einzelnen Kategorien und Diensten – und eine Suche anhand der aus der Spracheingabe des Nutzers extrahierten Suchworte nach passenden Knotenpunkten und/oder Dienstobjekten, welche diese Suchworte als Schlüsselworte aufweisen, kann der Dialog mit dem Benutzer relativ natürlich gestaltet werden. Der Benutzer muss bei der Suche den jeweiligen Dienst nicht entsprechend den vorgegebenen Kategorien klassifizieren, um zum Ziel zu gelangen, sondern er kann Formulierungen verwenden, die seines Erachtens den Dienst am besten beschreiben. Bei den Schlüsselworten handelt es sich daher vorzugsweise um den Namen des Dienstes bzw. der Kategorie selbst sowie um zusätzliche Schlüsselworte, wie insbesondere äquivalente Bezeichnungen des Dienstes bzw. der Kategorie oder Worte, die intuitiv mit diesem Dienst oder der Kategorie von den Nutzern in Verbindung gebracht werden können. Diese Vorgehensweise entspricht sogenannten „Shortcuts“ oder Abkürzungen in herkömmlichen Systemen mit dem Unterschied, dass diese nicht explizit erstellt und später aufwändig gepflegt werden müssen, sondern im Verfahren bereits „eingebaut“ sind. Durch die offene Struktur in der Weise, dass nicht nur ein Pfad entlang bestimmter festgelegter Knotenpunkte zu den Dienstobjekten führt, sondern dass die Datenstruktur in Form einer Mehrfach-Baumstruktur aufgebaut ist, wobei verschiedene Pfade entlang unterschiedlicher Knotenpunkte zum gleichen Dienstobjekt führen können, hat der Benutzer von mehreren Knotenpunkten aus die Möglichkeit, zum gleichen Dienstobjekt zu gelangen. Damit besteht die Möglichkeit, verschiedene Ordnungskriterien für ein Dienstobjekt festzulegen, die verschiedenen Benutzerkreisen mit unterschiedlichen Informations- und Wissensständen den bequemen Zugang zu den Diensten ermöglichen.
- Da die Schlüsselworte nur bis zu einem bestimmten vorgegebenen Akzeptanzkriterium mit den Suchworten übereinstimmen müssen, reicht es auch aus, wenn der Nutzer nicht wortgetreu alle Schlüsselworte einer Kategorie oder eines Dienstes als Suchworte in seiner Sprachausgabe nennt, sondern nur eine bestimmte Überdeckung zwischen Suchworten und Schlüsselworten vorliegt. Durch geschickte Wahl des Akzeptanzkriteriums kann dafür gesorgt werden, dass einerseits nicht zu viele Dienste oder

Kategorien gefunden werden, andererseits aber keine Kategorien oder Dienste verworfen werden, die möglicherweise zu dem vom Nutzer gewünschten Dienst führen oder bei denen es sich sogar um den gewünschten Dienst selbst handelt, deren Schlüsselworte jedoch aufgrund einer ungünstigen Spracheingabe des Nutzers nur eine Teilüberdeckung mit den Suchworten aufweisen. Das Akzeptanzkriterium sollte daher nicht zu eng gewählt werden.

Für die Suche der Knotenpunkte und Dienstobjekte innerhalb der Datenstruktur mittels der Suchworte kann beispielsweise ein Softwaremodul aus einer üblichen Internet-Suchmaschine verwendet werden, das die gefundenen Knotenpunkte und/oder Dienstobjekte – im Folgenden auch „Treffer“ genannt – mit einer prozentualen Trefferquote bewertet, die angibt, inwieweit eine Überdeckung zwischen Suchworten und Schlüsselworten vorliegt. Derartige Suchmodule sind dem Fachmann hinreichend bekannt und beispielsweise auch über die Produkte "FindIt" und "SpeechFinder" von Philips Speech Processing verfügbar. Die Suchmodule müssen lediglich von ihrer Datenschnittstelle an das Sprach-Dialogsystem angepasst sein bzw. umgekehrt. Es kann dann als Akzeptanzkriterium beispielsweise eine Trefferquote von 60 % angenommen werden.

Eine irrtümliche vorzeitige Verwerfung von eventuell richtigen, d. h. passenden Kategorien oder Diensten wird vermieden, da – sofern eine bestimmte Anzahl von Kandidaten-Knotenpunkten, d. h. eventuell passenden Kategorien, und/oder Kandidaten-Dienstobjekten, d. h. eventuell passenden Diensten, gefunden ist – alle diese Dienste und/oder Kategorien dem Benutzer vorzugsweise in Form einer bewerteten Liste angeboten werden. Dies erfolgt innerhalb eines Sprachausgabemenüs, welches von einer Prompt-Generierungseinheit erzeugt wird, d. h. es werden automatisch benutzerfreundliche vereindeutigende Fragen bzw. benutzerzugeschnittene Menüs in Abhängigkeit von den vorherigen Spracheingaben und dem Suchprozess kreiert, um dem Nutzer innerhalb des Dialogs zu helfen, die gewünschte Information zu finden bzw. den gewünschten Dienst zu erreichen. Durch die Unterteilung der Suche in mehrere Suchschritte, welche solange durchgeführt werden, bis nach einem Suchschritt die

Anzahl der ermittelten Kandidaten-Knotenpunkte und/oder Kandidaten-Dienstobjekte oberhalb einer vorgegebenen Minimalanzahl und unterhalb einer vorgegebenen Maximalanzahl liegt, wird dabei andererseits sichergestellt, dass dem Nutzer nicht zu lange Listen von Kategorien oder Diensten innerhalb eines Menüs des Dialogs angeboten werden. Die Maximalzahl sollte dementsprechend so gewählt werden, dass es sich um eine für den Nutzer akustisch leicht erfassbare und merkbare Anzahl von Kategorien oder Diensten handelt, so dass dieser nach dem Beenden der Menüausgabe noch alle angebotenen Dienste im Gedächtnis hat und entsprechend einen der Dienste oder eine der Kategorien auswählen kann.

10

Vorzugsweise sollte die Maximalanzahl gleich fünf gesetzt werden, so dass höchstens vier verschiedene Kategorien und/oder Dienste auf einmal angeboten werden.

Eine Möglichkeit, dies zu realisieren, besteht darin, dass das Akzeptanzkriterium variiert wird, wobei beispielsweise zunächst mit einem sehr großzügigen Akzeptanzkriterium gesucht wird und dann schrittweise, sofern zu viele Kandidaten-Knotenpunkte und/oder Kandidaten-Dienstobjekte ermittelt werden, das Akzeptanzkriterium verschärft wird, bis schließlich die Anzahl der Treffer, die das Akzeptanzkriterium erfüllen, innerhalb des gewünschten Bereichs liegt.

20

Bei einem besonders bevorzugten Ausführungsbeispiel werden die einem bestimmten Knotenpunkt zugeordneten Schlüsselworte automatisch auch den darunter klassifizierten weiteren Knotenpunkten oder Dienstobjekten zugeordnet, d. h. die Schlüsselworte werden innerhalb der Datenstruktur von oben nach unten weiter „vererbt“.

25

Es kann dann vorzugsweise nach einem erfolglosen Suchschritt – wobei „erfolglos“ hier so zu verstehen ist, dass entweder zu wenige oder zu viele Kandidaten-Knotenpunkte und/oder Kandidaten-Dienstobjekte gefunden wurden – die Suche auf einer anderen Ebene oder unter Einbeziehung einer anderen Ebene der Datenstruktur fortgeführt werden, bis die Anzahl der Kandidaten-Knotenpunkte und/oder Kandidaten-Dienst-

30

objekte innerhalb der gewünschten Grenzen liegt.

- Da durch die Vererbung sichergestellt ist, dass die Schlüsselworte von oben nach unten innerhalb der Datenstruktur ständig zunehmen und alle zu den höheren Kategorien, d. h. zu den weiter oben in der Baumstruktur befindlichen Knotenpunkten, gehörenden Schlüsselworte auch in einer unteren Ebene zu finden sind, wird vorzugsweise mit der Suche unten in der Datenstruktur, d. h. auf der Ebene der Dienstobjekte begonnen. Wird hier nicht das gewünschte Ergebnis erzielt, wird die Suche schrittweise, jeweils auf bzw. unter Einbeziehung einer nächsthöheren Ebene, unter den Knotenpunkten weitergeführt.
- 10 Bei diesem Verfahren ist es folglich nicht notwendig, das Akzeptanzkriterium selber zu verschärfen, sondern es kann einfach durch die schrittweise Suche auf verschiedenen Ebenen die Anzahl der Treffer reduziert werden, bis die Trefferanzahl innerhalb der gewünschten Grenzen liegt und ein sinnvolles Menü für die nächste Ausgabe an den Nutzer erzeugt werden kann. Dies hat den Vorteil, dass – anders als bei einer Ver-
- 15 schärfung des Akzeptanzkriteriums – keiner der in einem ersten Suchschritt gefundenen Treffer verworfen wird, da dies ja dazu führen könnte, dass gerade der richtige Treffer verworfen wird. Statt dessen wird ein Menü aus Kategorien einer höheren Ebene zusammengestellt, so dass gewährleistet ist, dass einerseits nur eine geringere Anzahl von Kategorien innerhalb des Menüs ausgegeben wird, aber die Kategorien andererseits
- 20 immer noch als „Oberbegriffe“ alle die Kategorien bzw. Dienste abdecken, die in dem vorhergehenden Suchschritt gefunden wurden.

- Mit anderen Worten wird bei diesem bevorzugten Ausführungsbeispiel des erfindungsgemäßen Dialogsystems dafür gesorgt, dass der Nutzer an einen Punkt innerhalb der Datenstruktur geführt wird, welcher sich einerseits möglichst nah an der untersten Ebene der Datenstruktur befindet, so dass vom Startpunkt des weiteren Dialogs aus nur wenige Abfragen notwendig sind, um das Dienstobjekt bzw. den Dienst zu erreichen. Andererseits befindet sich der Startpunkt in einer Ebene der Datenstruktur, die noch hoch genug ist, um alle anhand der extrahierten Suchwörter ermittelten Kategorien und Dienste
- 30 abzudecken und keine Treffer unnötig zu verwerfen.

- Vorzugsweise wird innerhalb eines Suchschritts bei einer zu geringen Anzahl von ermittelten Kandidaten-Knotenpunkten und/oder Kandidaten-Dienstobjekten das Akzeptanzkriterium erweitert. Dies ist insbesondere vorteilhaft bei dem zuletzt genannten Ausführungsbeispiel des Verfahrens, bei dem die Schlüsselworte von den
- 5 oberen Knotenpunkten der Datenstruktur an die darunter befindlichen Knotenpunkte und Dienste vererbt werden und bei denen mit der Suche schrittweise von unten nach oben begonnen wird, da hier bei dem ersten Suchschritt auf der Ebene der Dienste immer die meisten Treffer gefunden werden und ein weiterer Suchschritt auf einer höheren Ebene nur zu einer Verringerung der Trefferanzahl führen kann.
- 10 Bei Verwendung eines herkömmlichen Datenbank-Suchmoduls aus einer Internet-Suchmaschine kann eine Variation des Akzeptanzkriteriums einfach durch Veränderung der prozentualen Trefferquote realisiert werden.
- 15 Bei einem weiteren, relativ einfach realisierbaren Ausführungsbeispiel werden zur Suche die extrahierten Suchworte einzeln mit den Schlüsselworten jedes einzelnen Knotenpunkts und Dienstobjekts verglichen und die Anzahl der Übereinstimmungen zwischen Suchworten und Schlüsselworten für die einzelnen Knotenpunkte und Dienstobjekte gezählt. Das Akzeptanzkriterium kann dann einfach eine festgelegte Mindestanzahl von
- 20 Übereinstimmungen zwischen den extrahierten Suchworten und den Schlüsselworten sein. So kann beispielsweise festgelegt werden, dass alle Suchworte innerhalb der Schlüsselworte eines Dienstobjekts oder eines Knotenpunkts enthalten sein müssen oder mindestens zwei Suchworte oder mindestens ein Suchwort etc.
- 25 Anspruch 10 beschreibt eine weitere besonders vorteilhafte Variante des erfindungsgemäßen Sprach-Dialogsystems. Sie betrifft den Fall, dass der Nutzer nach Durchführen einer Suche und einer Ansage eines Menüs gemäß dem erfindungsgemäßen Verfahren eine Spracheingabe tätigt, die weitere, neue Suchworte enthält.

30

Ein Beispiel hierfür ist der Fall, dass der Benutzer zunächst mit dem Suchwort „Reise“ folgenden Prompt erhält: „Möchten Sie mit dem Auto, mit der Bahn oder mit dem Flugzeug reisen? Bitte wählen Sie einen Begriff aus“, und der Benutzer dann relativ salopp antwortet: „Auto ist gut, ich will mobil sein.“ Dieser Antwortsatz des Benutzers

5 enthält zwei potentielle Schlüsselworte, nämlich die Begriffe „Auto“ und „mobil“. Die anderen Worte dieses Satzes werden bei der Analyse als nicht sinntragend erkannt. Dementsprechend werden aus dieser Spracheingabe zwei neue Suchworte, nämlich „Auto“ und „mobil“ extrahiert. Das Sprach-Dialogsystem führt dann eine erneute Suche mit den Suchworten „Auto“ und „mobil“ durch und findet beispielsweise die Kategorie

10 „Auto“ (wie bei der ersten Suche) und zusätzlich die Kategorie „Mobilfunkgerät“, welche zu bestimmten Telefonauskunftsdiensten oder Tarifauskunftsdiensten führen kann. Wird nun das Ergebnis der ersten Suche, bestehend aus den Kategorien „Auto, Bahn, Flugzeug“, mit dem Ergebnis der zweiten Suche, bestehend aus den Kategorien „Auto“ und „Mobilfunk“, geschnitten, so erhält man als Gesamtergebnis die Kategorie

15 „Auto“, bei der es sich eindeutig um die vom Nutzer gewünschte Kategorie handelt. Diese Kategorie wird dann bevorzugt an den Nutzer wieder ausgegeben.

Sofern die Schnittmenge mehrere Treffer enthält, kann der Nutzer dann unter diesen bevorzugten Treffern auswählen. Bei nur einem einzelnen Schnittmengenelement kann

20 die bevorzugte Ausgabe nur für eine Verifikation durch den Nutzer erfolgen, beispielsweise durch die Ansage „Sie haben 'Auto' gewählt, ist das richtig?“

Sofern die Schnittmenge leer ist, wird vom Sprach-Dialogsystem das vorherige Suchergebnis ignoriert und nur das neue Suchergebnis genutzt. Ein Beispiel hierfür wäre, wenn

25 der Nutzer auf die erste Ausgabe des ersten Suchergebnisses antwortet: „Ich wollte eigentlich Informationen über Mobilfunktarife haben.“ Dieser Antwortsatz des Benutzers enthält lediglich das Suchwort „mobil“ und führt zu einem Suchergebnis, welches nur die Kategorie „Mobilfunk“ enthält. Die Schnittmenge zwischen dem ersten Suchergebnis „Auto, Bahn, Flugzeug“ und dem zweiten Suchergebnis „Mobilfunk“ ist folglich leer,

30 und es wird zu Recht nach Wunsch des Nutzers lediglich die Kategorie „Mobilfunk“ angeboten.

Die Erfindung wird im Folgenden unter Hinweis auf die beigefügten Figuren anhand eines Ausführungsbeispiels näher erläutert. Es zeigen:

5      **Figur 1 ein schematisches Blockschaltbild der wesentlichen Komponenten eines erfindungsgemäßen Sprach-Dialogsystems,**

10      **Figur 2 eine Blockdarstellung eines einfachen, anschaulichen Beispiels für eine Datenstruktur in einer Datenbank eines erfindungsgemäßen Sprach-Dialogsystems,**

15      **Figur 3 einen Teil eines Flussdiagramms für einen möglichen Verfahrensablauf bei der Nutzung des Sprach-Dialogsystems.**

Bei dem in Figur 1 dargestellten Ausführungsbeispiel handelt es sich um ein Sprach-Dialogsystem 1, welches über eine Netzanschlussschnittstelle 5 verfügt, über die das Sprach-Dialogsystem 1 an ein öffentliches Kommunikationsnetz, beispielsweise an ein Telefonnetz, angeschlossen ist und somit von einem Nutzer über ein normales Telefon 14 erreichbar ist.

20      Um eine Kommunikation mit dem Nutzer in natürlicher Sprache durchzuführen, weist das Sprach-Dialogsystem 1 eine Spracherkennungseinheit 2 auf. Diese empfängt die von der Netzanschlussschnittstelle 5 ankommenden Sprachsignale des Nutzers und führt eine Spracherkennung durch, wobei die innerhalb des Sprachsignals enthaltenen Informationen in Daten umgesetzt werden, die von den nachfolgenden Teilen des Systems verarbeitet werden können. Ausgangsseitig weist das Sprach-Dialogsystem 1 eine Spracherzeugungseinrichtung 3 auf. Hierbei kann es sich beispielsweise um ein sogenanntes TTS-System (Text-To-Speech-System) handeln, welches aus eingehenden, computerlesbaren Daten durch eine passende Zusammensetzung von Phonemen und Wörtern einen gesprochenen Text erzeugt. Es kann sich aber auch um einen sogenannten  
30      **Prompt-Player handeln, welcher fertig abgespeicherte Texte enthält, die entsprechend**

aufgerufen und dem Nutzer abgespielt werden. Ebenso kann es sich auch um eine Einrichtung handeln, welche kombiniert ein TTS-System und einen Prompt-Player nutzt. Die ausgehenden Sprachdaten werden wiederum über die Netzanschlussschnittstelle 5 an das Telefon 14 des Nutzers übermittelt.

5

Kernstück des Sprach-Dialogsystems 1 ist eine Dialogkontrolleinrichtung 4, welche gemeinsam mit einer Datenbank 6 den Dialog mit dem Benutzer steuert und die im Dialogsystem 1 enthaltenen Dienste 9 aufruft oder den Nutzer über eine Call-Transfer-Einheit 7 zu einem externen Dienst 10 vermittelt.

10

Das dargestellte Sprach-Dialogsystem 1 kann im Wesentlichen in Form von Software auf einem geeigneten Rechner bzw. Server realisiert sein. Hierbei können die Spracherkennungseinrichtung 2, die Spracherzeugungseinrichtung 3 und die Dialogkontrolleinrichtung 4 reine Softwaremodule sein, die auf geeignete Weise untereinander verknüpft sind. Lediglich die Netzanschlussschnittstelle 5 muss entsprechende Hardwarekomponenten zum Anschluss an das gewünschte Netz aufweisen. Da auch ein Call-Transfer hardwaremäßig über die Netzanschlussschnittstelle 5 durchgeführt werden kann, kann – anders als in Figur 1 dargestellt – im Prinzip auch die Call-Transfer-Einheit 7 eine reine Softwareeinheit sein, die nur die notwendigen Informationen zur Durchführung eines Call-Transfers zu den verschiedenen externen Diensten 10 enthält und über die Netzanschlussschnittstelle 5 den Call-Transfer innerhalb des Kommunikationsnetzes einleitet.

15

20

Neben den in Figur 1 dargestellten Komponenten kann das Sprach-Dialogsystem auch noch weitere Komponenten aufweisen, wie sie üblicherweise in Sprach-Dialogsystemen verwendet werden. Als Beispiel ist hier eine zusätzliche Datenbank 8 gezeigt, welche verschiedene Informationen über einzelne Nutzer enthält, die gegenüber dem Sprach-Dialogsystem registriert sind und die bei einem Anruf identifiziert werden können. Eine solche Datenbank kann insbesondere Informationen über von den Nutzern bevorzugt genutzten Dienste, über einen letzten Gebrauch des Sprach-Dialogsystems durch den betreffenden Nutzer oder Ähnliches enthalten. Die zusätzlichen Informationen können

25

30



dann bei der Spracherkennung, bei der Analyse von Suchworten oder auf sonstige Weise für den Nutzer verwendet werden, um ihn so schneller zu den gewünschten Diensten zu geleiten. Insbesondere kann das Sprach-Dialogsystem auch noch zusätzliche Komponenten zur Führung einer Statistik über die Nutzung des Sprach-Dialogsystems oder einzelner Dienste allgemein oder für spezielle Nutzer enthalten.

Die Dialogkontrolleinrichtung 4 selbst besteht in dem dargestellten Ausführungsbeispiel aus mehreren, passend verknüpften Software-Modulen.

- 10 Hierzu zählt insbesondere ein Analysemodul 11, welches aus den von der Spracherkennungseinrichtung 2 empfangenen Daten bestimmte Suchworte extrahiert. Diese Suchwort-Extraktion erfolgt auf Basis von vorgegebenen grammatikalischen und syntaktischen Regeln, damit nicht jedes Wort innerhalb eines vom Nutzer gesprochenen Satzes als Suchwort extrahiert wird und insbesondere die nicht sinntragenden Worte
- 15 innerhalb des Satzes ignoriert werden. So werden beispielsweise aus dem Satz „Ich möchte Theaterinformationen“ die Begriffe „Theater“ und „Informationen“ als Suchworte extrahiert, wogegen die Worte „ich möchte“ keine weitere Bedeutung für die weitere Verarbeitung haben.
- 20 Anhand der Suchworte wird dann von einem Suchmodul 12 der Dialogkontrolleinrichtung 4 in der Datenbank 6 eine Suche nach bestimmten Diensten und/oder Kategorien durchgeführt. Diese Datenbank enthält eine Datenstruktur DS nach Art eines Mehrfach-Entscheidungsbaums. Ein Beispiel hierfür ist in Figur 2 dargestellt. Die Datenstruktur DS enthält hier mehrere Knotenpunkte K, welche über Pfade P miteinander verbunden
- 25 sind. Die Knotenpunkte K befinden sich in zwei Ebenen I, II. Auf einer dritten Ebene III unter der unteren Knotenpunktebene II befindet sich eine Ebene mit Dienstobjekten D.

Diese Dienstobjekte D repräsentieren die einzelnen Dienste 9, 10. Bei dem dargestellten Ausführungsbeispiel handelt es sich jeweils um komplexere Dienste, bei denen noch

30 weitere Abfragen innerhalb des Dienstes notwendig sind, damit der Benutzer zu der

gewünschten Information kommt. So kann es sich beispielsweise bei dem Dienst „Festnetz“ um eine normale Telefonauskunft handeln, an welche der Nutzer weitervermittelt wird. Bei dem Dienst „Bahn“ handelt es sich um einen Dienst einer Eisenbahngesellschaft, zu welchem der Nutzer weitergeleitet wird, sobald er den Dienst „Bahn“ auswählt. Die einzelnen Dienste 9, 10 können hierbei ebenfalls wieder in der erfindungsgemäßen Weise als Sprach-Dialogsystem aufgebaut sein. So kann beispielsweise eine Telefonauskunft, welche sich hinter dem Dienst „Festnetz“ verbirgt, wiederum eine Datenbank mit einer eigenen Baumstruktur mit mehreren Kategorien und Diensten aufweisen, wobei als Dienst dann letztendlich die Ausgabe einer gesuchten Information über einen bestimmten Teilnehmer, z. B. die Telefonnummer oder die Adresse, zu verstehen ist.

Die Knotenpunkte K innerhalb der Datenstruktur DS repräsentieren jeweils Kategorien, in welche die in der Ebene darunter befindlichen Kategorien bzw. Unterkategorien oder Dienste klassifizierbar bzw. einzuordnen sind. Wie in Figur 2 deutlich zu sehen ist, ist jeder der Dienste zumindest in einer Kategorie in der mittleren Ebene II eingeordnet. Hierbei fallen mehrere Dienste ggf. unter die gleiche Kategorie, so dass umgekehrt von einem Knotenpunkt K der mittleren Ebene II mehrere Pfade P zu verschiedenen Diensten D führen können. Ebenso sind die Kategorien der mittleren Ebene II als Unterkategorien den Kategorien in der Ebene I zugeordnet.

In Figur 2 ist aus Gründen der Anschaulichkeit nur ein sehr einfaches Ausführungsbeispiel einer erfindungsgemäßen Datenstruktur DS dargestellt. In der Realität ist eine solche Datenstruktur erheblich komplexer und erstreckt sich über eine Vielzahl von Ebenen mit jeweils einer Vielzahl von parallelen Knotenpunkten und/oder Dienstobjekten. Dabei muss im Übrigen nicht jeder Dienst oder Knotenpunkt einer Kategorie der nächsthöheren Ebene zugeordnet sein, sondern es können auch eine oder mehrere Ebenen durch einen Pfad übersprungen werden.

Jedem der einzelnen Knotenpunkte K und Dienstobjekte D sind verschiedene Schlüsselworte S zugeordnet. Zu diesen Schlüsselworten S zählen insbesondere die Namen der einzelnen Kategorien bzw. Dienste selbst, wie sie in den Kästchen in Figur 2 genannt sind. Darüber hinaus können den einzelnen Knotenpunkten K und Dienstobjekten D –  
5 bzw. den Kategorien und Diensten – jeweils zusätzliche Schlüsselworte zugeordnet sein. Sinnvollerweise gehören zu den weiteren Schlüsselworten geeignete Synonyme der Namen der einzelnen Dienste oder Kategorien oder sonstige Schlüsselworte, unter denen der Nutzer üblicherweise einen solchen Dienst oder eine Kategorie suchen würde oder die mit dem Dienst in Beziehung stehen können. Beispielsweise können, wie in Figur 2  
10 dargestellt, dem Dienst „Auto“ die Schlüsselworte „Abfahrt“ und „Wetter“ zugeordnet sein, dem Dienst „Bahn“ die Schlüsselworte „Abfahrt“ und „Ankunft“ und dem Dienstobjekt „Flug“ die Schlüsselworte „Reiseland“ und „Wetter“.

Die Schlüsselworte einer Kategorie werden an die zugehörigen Kategorien oder Dienste  
15 in der nächsttieferen Ebene „weitervererbt“. Dies ist in Figur 2 beispielhaft an der Kette der Kategorie „Ort“, Unterkategorie „mobil“ und Dienst „Auto“ dargestellt. Zu der Kategorie „Ort“ gehören die Schlüsselworte „Ort“ und „Standpunkt“, zu der Unterkategorie „mobil“ gehören dann die Schlüsselworte „mobil“, „Ort“, „Standpunkt“ sowie  
20 eventuell weitere Schlüsselworte, und mit dem Dienst „Auto“ sind dann die Schlüsselworte „Auto“, „Ort“, „Standpunkt“, „Info“, „mobil“, „Reise“ sowie ggf. weitere Schlüsselworte verknüpft.

Die Suche in dieser Datenstruktur DS verläuft in dem dargestellten Beispiel mit drei Ebenen auf folgende Weise:

25 Begonnen wird mit einem Suchschritt auf der untersten Ebene III, d. h. auf der Ebene III der Dienstobjekte D. Anhand der Überdeckung der Suchworte mit den Schlüsselworten S der einzelnen Dienstobjekte D werden hier Kandidaten-Dienstobjekte D gesucht, welche zu dem gewünschten Dienst des Benutzers passen könnten. Für die konkrete  
30 Durchführung der Suche gibt es dabei verschiedene Möglichkeiten. Zum einen kann ein



durch einen Tastendruck des Telefons, beispielsweise mittels eines DTMF-Verfahrens, erfolgen kann. Hierzu kann beispielsweise bei der Prompt-Generierung vor dem jeweiligen Treffer, d. h. dem gefundenen Dienst oder der Kategorie, eine Zahl angesagt werden, und der Nutzer kann dementsprechend nur die passende Taste seiner Zifferntastatur am Telefon drücken. Das Sprach-Dialogsystem muss hierfür natürlich zusätzlich  
5 entsprechende Mittel zur Erkennung und Verarbeitung der DTMF-Signale aufweisen.

Überschreitet die Anzahl der Treffer die vorgegebene Maximalanzahl, so wird ein erneuter Suchschritt durchgeführt. Hierbei wird die Suche auf der nächsthöheren Ebene  
10 – in dem dargestellten Beispiel auf der mittleren Ebene II, d. h. der Ebene direkt über den Diensten angeordneten Knotenpunkten K – fortgesetzt. Da mehrere Dienste in eine Kategorie fallen und die Schlüsselworte nach unten vererbt werden, ist die Anzahl der Kategorien auf dieser Ebene geringer als die Anzahl der Dienste auf der darunter befindlichen Ebene III. Somit ist bei einer Suche mit den gleichen Suchworten die  
15 Anzahl der Treffer auf dieser Ebene II geringer als bei dem vorhergehenden Suchschritt auf der darunter befindlichen Ebene III.

In dem in Figur 2 dargestellten Beispiel liegt naturgemäß, da die Ebene II nur vier Kategorien aufweist, die Anzahl der Treffer, d. h. der möglichen Kandidaten-Knotenpunkte,  
20 immer unter oder gleich vier. In der Realität wird auch diese Ebene erheblich mehr als vier unterschiedliche Kategorien bzw. Knotenpunkte aufweisen, so dass in vielen Fällen auch auf dieser Ebene die Anzahl der gefundenen Kandidaten-Knotenpunkte noch die Maximalanzahl überschreitet. In diesem Fall wird auf der nächsthöheren Ebene solange weitergesucht, bis schließlich die Anzahl der gefundenen Kandidaten-Knotenpunkte bzw.  
25 der evtl. passenden Kategorien unter der Maximalanzahl liegt.

Bei der Suche muss natürlich berücksichtigt werden, wenn nicht jeder Dienst bzw. Kategorie einer Kategorie der nächsthöheren Ebene zugeordnet ist, sondern eine oder mehrere Ebenen durch einen Pfad übersprungen werden. Dann müssen gegebenenfalls  
30 bei einem neuen Suchschritt in der nächst höheren Ebene die im vorhergehenden

Suchschritt bereits gefundenen Kandidaten-Dienstobjekte oder Kandidaten-Knotenpunkte, die nicht mit einem Knotenpunkt der höheren Ebene verbunden sind, wieder in die Suche mit einbezogen werden.

- 5 Wird bei einem Suchschritt schließlich nur noch ein Kandidaten-Knotenpunkt ermittelt, so wird die Suche abgebrochen, da hier eine Situation erreicht ist, in der es nicht mehr möglich ist, die Treffer-Anzahl noch weiter zu reduzieren. Ein Fortsetzen der Suche auf einer tieferen Ebene ist ebenfalls nicht sinnvoll, da die Suche hier ja bereits durchgeführt wurde und zu einem Ergebnis mit einer Trefferanzahl geführt hat, welche höher ist als
- 10 die vorgegebene Maximalanzahl. Dies bedeutet, dass alle diese Treffer zu dem einen gefundenen Kandidaten-Knotenpunkt der höheren Ebene gehören. Ein Beispiel für eine solche Situation ist eine Abfrage nach einer Verkaufsstelle eines Unternehmens, wobei die Verkaufsstellen jeweils nach Verkaufsgebieten in verschiedenen Postleitzahlregionen sortiert sind und die Anzahl der Verkaufsgebiete die Anzahl der Maximaltreffer über-
- 15 schreitet. In einer solchen Situation werden daher unabhängig von der jeweiligen Anzahl alle unter den jeweiligen zuletzt gefundenen Kandidaten-Knotenpunkt gehörigen Kategorien oder Dienste an den Nutzer innerhalb eines Sprachausgabemenüs ausgegeben.

- Da in diesem Fall die Maximalanzahl überschritten ist, erfolgt die Ausgabe vorzugsweise
- 20 gruppiert und unter Angabe einer Gruppenbezeichnung. Bei dieser Gruppenbezeichnung handelt es sich beispielsweise um eine Zahl oder einen Namen, so dass der Benutzer zunächst eine Gruppe durch Angabe der Gruppenbezeichnung auswählen kann und dann diese Gruppe von Kategorien oder Diensten noch einmal zur weiteren Auswahl angesagt bekommt. Alternativ ist es auch möglich, dass das Sprach-Dialogsystem von selbst
- 25 zunächst eine vereindeutigende Frage generiert und anhand der Antwort des Benutzers dann eine der beiden Gruppen zur Ansage auswählt. Bei dem genannten Beispiel könnte das Sprach-Dialogsystem nach dem Wohnort des Nutzers fragen und dann zunächst nur die Verkaufsgebiete der näheren Umgebung zur Auswahl anbieten. Wird bei einem Suchschritt in einer höheren Ebene kein Kandidatenpunkt ermittelt, so wird entsprechend
- 30 die vollständige Liste der zuletzt ermittelten Kategorien oder Dienste ausgegeben.

Figur 3 zeigt einen Ausschnitt aus einem Flussdiagramm, welches den möglichen Ablauf eines Dialogs bei der Nutzung des Sprach-Dialogsystems 1 darstellt. Nach Eingabe eines Sprachbefehls durch den Nutzer erfolgt zunächst die Spracherkennung. Dann werden aus den erkannten Sprachinformationen die Suchworte extrahiert. Anschließend erfolgt anhand dieser Suchworte eine Suche gemäß dem vorbeschriebenen Verfahren. Ist genau ein Dienst gefunden, so wird der betreffende Dienst für den Nutzer aufgerufen bzw. wenn es sich bei dem Dienst um eine reine Informationsausgabe handelt, diese Information ausgegeben. Ansonsten wird zunächst ein Prompt generiert und ausgegeben, mit dessen Hilfe der Nutzer zur Auswahl einer Kategorie oder eines Dienstes aus einer Anzahl von Kandidaten-Kategorien oder Kandidaten-Diensten aufgefordert wird. Die darauf vom Nutzer erfolgende Antwort wird dann wiederum der Spracherkennung zugeführt und eine erneute Suchwortextrahierung veranlasst. Mit den neuen Suchworten wird dann die Suche fortgesetzt. Dieses Verfahren wird so lange fortgeführt, bis schließlich der gewünschte Dienst gefunden ist, oder ein expliziter Abbruch des Dialogs, beispielsweise auf Wunsch des Nutzers, erfolgt.

Im Folgenden wird außerdem eine Alternative zur Verwendung einer handelsüblichen Internet-Suchmaschine bzw. eines Suchmoduls einer solchen Suchmaschine beschrieben. Bei diesem Ausführungsbeispiel wird die Datenbank jeweils nach jedem Suchwort einzeln durchsucht und für jedes Suchwort als Ergebnis eine Menge von Knotenpunkten oder Dienstobjekten ermittelt, deren Schlüsselworte dieses bestimmte Suchwort enthalten. Die Anzahl der Übereinstimmungen der Suchworte mit den Schlüsselworten wird als Akzeptanzkriterium verwendet. Dies ist relativ einfach durch eine geeignete Bildung von Schnittmengen und/oder Vereinigungsmengen der Suchergebnisse möglich.

Das engste Akzeptanzkriterium ist bei diesem Verfahren die Festlegung, dass nur solche Treffer akzeptiert werden, bei denen alle Suchworte innerhalb der Schlüsselworte identisch vorhanden sind. Diejenigen Kategorien bzw. Dienste, deren Schlüsselworte alle Suchworte enthalten, lassen sich durch eine Schnittmengenbildung gemäß folgender Regel ermitteln:

$$\bigcap_{i \in \{1, \dots, n\}} A_i \quad (1)$$

Die  $A_i$  stellen dabei jeweils das Suchergebnis für das  $i$ -te Suchwort dar, d. h. die Menge der Kategorien bzw. Dienste, deren Schlüsselworte das  $i$ -te Suchwort enthalten.

5 Gemäß der Regel

$$\bigcup_{i \neq j} (A_i \cap A_j) \quad (2)$$

lassen sich alle die Kategorien bzw. Dienste herausfinden, welche wenigstens zwei der  
10 Suchworte unter ihren Schlüsselworten aufweisen.

Außerdem können gemäß der Regel

$$\bigcup_{i \in \{1, \dots, n\}} A_i \quad (3)$$

15

alle Kategorien bzw. Dienste ermittelt werden, welche wenigstens ein Schlüsselwort aufweisen, das mit einem der Suchworte übereinstimmt.

Es stehen somit drei unterschiedlich strenge Akzeptanzkriterien zur Verfügung. Daher  
20 kann zunächst in einem Suchschritt die Anzahl der Treffer nach Formel (1) ermittelt werden. Wenn die Trefferanzahl zu gering ist, erfolgt eine Berechnung nach Formel (2) und schließlich nach Formel (3). Wird letztlich beim dritten Akzeptanzkriterium immer noch kein Treffer gefunden, so wird der Benutzer zur Eingabe einer neuen Suchanfrage aufgefordert.

25

Die Formeln (1) bis (3) werden im Folgenden noch mal an einem konkreten Beispiel verdeutlicht. Hierzu wird auf Figur 2 verwiesen.



Es sei angenommen, dass dem Dienst „Auto“ die Suchworte „Abfahrt“ und „Wetter“ zugeordnet sind, dem Dienst „Bahn“ die Suchworte „Abfahrt“ und „Ankunft“ und dem Dienst „Flug“ die Suchworte „Reiseland“ und „Wetter“.

- 5 Wird dann weiterhin davon ausgegangen, dass aus der Spracheingabe des Benutzers das Suchwort „Abfahrt“ ermittelt wurde, so ist das Suchergebnis bei diesem einen Suchwort  $A_1 = \{\text{„Auto“}, \text{„Bahn“}\}$ . Da nur ein Suchwort vorhanden ist und somit nur eine Treffermenge  $A_1$  existiert, ändert sich durch die Schnittmengenbildung gemäß Formel (1) nichts.

10

Anders ist dies, wenn zusätzlich nach einem zweiten Suchwort, hier beispielsweise dem Suchwort „Wetter“ gesucht wird. Das Suchergebnis für dieses zweite Suchwort „Wetter“ wäre dann  $A_2 = \{\text{„Flug“}, \text{„Auto“}\}$ . Die Bildung der Schnittmenge von  $A_1$  und  $A_2$  führt zu

15

$$A_1 \cap A_2 = \{\text{„Auto“}\} \quad (4)$$

20

d. h. nur der Dienst „Auto“ enthält sowohl das Suchwort „Abfahrt“ als auch das Suchwort „Wetter“ als Schlüsselworte. Somit wird exakt ein Dienst gefunden, welcher das strengste Akzeptanzkriterium erfüllt, und der Nutzer wird an diesen Dienst vermittelt.

25

Ein anderer Fall liegt vor, wenn noch ein drittes Schlüsselwort, beispielsweise das Schlüsselwort „Ankunft“ hinzukommt. Man erhält dann für das Suchwort „Abfahrt“ das Suchergebnis  $A_1 = \{\text{„Auto“}, \text{„Bahn“}\}$ , für das Suchwort „Wetter“ das Suchergebnis  $A_2 = \{\text{„Auto“}, \text{„Flug“}\}$  und für das dritte Suchwort „Ankunft“ das Suchergebnis  $A_3 = \{\text{„Bahn“}\}$ . Erfolgt die Ermittlung des gesamten Suchergebnisses nach Formel (1), so erhält man eine leere Menge, da keine der Dienste alle Suchbegriffe innerhalb der Schlüsselworte enthält. Wird das Akzeptanzkriterium reduziert und die Berechnung des Gesamtergebnisses gemäß Formel (2) durchgeführt, so erhält man

30

$$A_1 \cap A_2 = \{\text{"Auto"}\}$$

$$A_1 \cap A_3 = \{\text{"Bahn"}\}$$

$$A_2 \cap A_3 = \emptyset$$

$$\Rightarrow (A_1 \cap A_2) \cup (A_1 \cap A_3) \cup (A_2 \cap A_3) = \{\text{"Auto"}, \text{"Bahn"}\} \quad (5)$$

5

Man erhält folglich die Dienste „Auto“ und „Bahn“ als Gesamtergebnis, da diese beiden jeweils zwei der Suchworte innerhalb ihres Schlüsselwortsatzes enthalten.

Werden dagegen als Suchworte die Begriffe „Reiseland“ und „Ankunft“ gewählt, so erhält man für das erste Suchwort „Reiseland“ das Suchergebnis  $A_1 = \{\text{"Flug"}\}$  und für das zweite Suchwort „Ankunft“ das Ergebnis  $A_2 = \{\text{"Bahn"}\}$ . Wird nun das engste Akzeptanzkriterium gewählt und eine Berechnung des Gesamtergebnisses gemäß Formel (1) durchgeführt, so erhält man eine leere Menge. Ebenso führt eine Berechnung gemäß Formel (2) zu einer leeren Menge, weil in keinem der gefundenen Treffer beide Suchworte innerhalb der Schlüsselwortlisten vorhanden sind. Erst eine Berechnung des Gesamtergebnisses nach Formel (3) führt dazu, dass alle gefundenen Treffer, nämlich die Dienste „Bahn“ und „Flug“ als Kandidaten-Dienste bzw. Kandidaten-Dienstobjekte akzeptiert werden.

Es wird noch einmal darauf hingewiesen, dass neben den konkret beschriebenen Suchverfahren bzw. der Verwendung eines handelsüblichen Suchmoduls auch beliebige weitere Suchalgorithmen verwendet werden können, um das erfindungsgemäße Verfahren durchzuführen. Ebenso kann das Verfahren auch an weiteren Stellen modifiziert werden, ohne die Erfindung wesentlich zu verändern. So ist beispielsweise zu beliebigen Zeitpunkten jederzeit eine Ausgabe weiterer Prompts zur zusätzlichen Verifikation von Zwischenergebnissen möglich.

Das System kann auch als sogenanntes „Barge-In“-Dialogsystem aufgebaut sein, bei dem der Nutzer jederzeit während der Ausgabe eines Prompts antworten kann und diese Antwort vom Sprach-Dialogsystem akzeptiert und bearbeitet wird und die weitere

Ausgabe des Prompts unterbrochen wird. Ebenso kann eine Suche jederzeit auf Wunsch des Nutzers oder bei Eintritt eines zuvor festgelegten Abbruchkriteriums abgebrochen werden.

- 5 Weiterhin wird noch einmal darauf hingewiesen, dass es sich bei dem in Figur 1 dargestellten Ausführungsbeispiel nur um eine sehr grob vereinfachte Darstellung des Sprach-Dialogsystems handelt und dass das erfindungsgemäße Sprach-Dialogsystem auch in modifizierter Form realisiert werden kann. Insbesondere ist es möglich, die einzelnen Softwaremodule anstatt auf einem einzelnen Rechner auf verschiedenen  
10 Rechnern innerhalb eines Netzes anzuordnen, wobei es sich insbesondere anbietet, besonders rechenintensive Funktionen wie beispielsweise die Spracherkennung auf andere Rechner auszulagern. Weiterhin ist es möglich, dass das Sprach-Dialogsystem alternativ oder zusätzlich zum Telefonanschluss eine eigene Benutzerschnittstelle mit einem Mikrofon und einem Lautsprecher aufweist. Auch das Bereitstellen von  
15 Sprachdaten über das Datennetz – sog. Voice-over-IP – ist möglich. Mit Hilfe der Erfindung ist es möglich, ein Voice-Portal aufzubauen, welches erheblich intuitiver und flexibler vom Nutzer zu bedienen ist als bisher bekannte Voice-Portale. Außerdem ist ein solches Sprach-Dialogsystem in der Lage, große Datenbanken, beispielsweise für telefonische Branchenauskunftssysteme bzw. sog. Yellow-Page-  
20 Applikationen, zu verwalten. Zudem können die Nutzer in einem relativ einfachen und effizienten Verfahren ihre Suchanfrage formulieren und verfeinern. Dabei wird eine Ausgabe von langen, unhandlichen und für den Benutzer nicht mehr merkbaren Listen vermieden.

### PATENTANSPRÜCHE

1. Verfahren zum Betrieb eines Sprach-Dialogsystems (1), welches mit einem Nutzer unter Verwendung einer Spracherkennungseinrichtung (2) und einer Sprachausgabereinrichtung (3) kommuniziert,
- wobei dem Nutzer verschiedene Dienste (9, 10) innerhalb des Sprach-Dialogsystems (1) oder durch Vermittlung des Sprach-Dialogsystems (1) zur Verfügung stehen und vom Nutzer innerhalb eines mit dem Sprach-Dialogsystem (1) geführten Dialogs auswählbar sind,
- und dabei zur Steuerung des Dialogs zum Auswählen eines Dienstes (9, 10) durch den Nutzer eine Datenbank (6) verwendet wird, mit einer hierarchischen Datenstruktur (DS) mit einer Mehrzahl von Knotenpunkten (K) und einer Mehrzahl von Pfaden (P) zur Verbindung der Knotenpunkte (K) untereinander und zur Verbindung von Knotenpunkten (K) mit Dienstobjekten (D), welche jeweils an einem Ende eines Pfades (P) in der Datenstruktur (DS) angeordnet sind, wobei die Dienstobjekte (D) die zur Verfügung stehen Dienste (9, 10) repräsentieren und die Knotenpunkte (K) jeweils Kategorien repräsentieren, in welche wiederum andere Kategorien und/oder Dienste, die durch die innerhalb der hierarchischen Datenstruktur (DS) auf einer Ebene (II, III) unter dem jeweiligen Knotenpunkt (K) angeordneten weiteren Knotenpunkte (K) oder Dienstobjekte (D) repräsentiert werden, klassifiziert sind,
- dadurch gekennzeichnet,
- 20 dass zumindest zu einem Teil der Dienstobjekte (D) und/oder Knotenpunkte (K) innerhalb der Datenstruktur (DS) eine Mehrzahl von Pfaden (P) führt, und jedem Knotenpunkt (K) und jedem Dienstobjekt (D) ein oder mehrere Schlüsselworte (S) zugeordnet sind

15

20

25

## 4. Verfahren nach Anspruch 3,

dadurch gekennzeichnet,

- 5 dass mit der Suche innerhalb der Datenstruktur (DS) auf der Ebene (I) der Dienstobjekte (D) begonnen wird und dann die Suche schrittweise jeweils auf oder unter Einbeziehung einer nächst höheren Ebene (II, III) unter den Knotenpunkten (K) weitergeführt wird.

## 5. Verfahren nach einem der Ansprüche 1 bis 4,

dadurch gekennzeichnet,

- 10 dass die vorgegebene Minimalanzahl der Kandidaten-Knotenpunkte (K) und/oder Kandidaten-Dienstobjekte (D) gleich eins ist und bei Ermittlung nur eines Kandidaten-Dienstobjekts (D) in einem Suchschritt der durch dieses Kandidaten-Dienstobjekt (D) repräsentierte Dienst (9, 10) aufgerufen wird.

## 6. Verfahren nach einem der Ansprüche 1 bis 5,

15 dadurch gekennzeichnet,

- dass die vorgegebene Minimalanzahl der Kandidaten-Knotenpunkte (K) und/oder Kandidaten-Dienstobjekte (D) gleich eins ist und bei Ermittlung nur eines Kandidaten-Knotenpunkts (K) in einem Suchschritt die Suche abgebrochen wird und unabhängig von der jeweiligen Anzahl aller Kategorien oder Dienste (9,10) an den Nutzer in einem
- 20 Sprachausgabe-Menü ausgegeben werden, die durch die Knotenpunkte (K) und/oder Dienstobjekte (D) repräsentiert werden, die sich in der Datenstruktur (DS) unter dem betreffenden Kandidaten-Knotenpunkt (K) befinden und mit diesem über einen Pfad (P) verbunden sind.

25 7. Verfahren nach Anspruch 6,

dadurch gekennzeichnet,

- dass die Kategorien oder Dienste gruppiert unter Angabe einer Gruppenbezeichnung ausgegeben werden.

8. Verfahren einem der Ansprüche 1 bis 7,

dadurch gekennzeichnet,

dass das Akzeptanzkriterium eine Mindestanzahl von Übereinstimmungen zwischen den extrahierten Suchworten und den einem Knotenpunkt oder Dienstobjekt zugeordneten

5 Schlüsselworten umfasst.

9. Verfahren nach einem der Ansprüche 1 bis 8,

dadurch gekennzeichnet,

dass innerhalb eines Suchschritts bei einer zu geringen Anzahl von ermittelten

10 Kandidaten-Knotenpunkten und/oder Kandidaten-Dienstobjekten das Akzeptanzkriterium erweitert wird.

10. Verfahren nach einem der Ansprüche 1 bis 9,

dadurch gekennzeichnet,

15 dass, sofern das Sprach-Dialogsystem nach einer ersten Suche und einer Ansage der durch die ermittelten Kandidaten-Knotenpunkte und/oder Kandidaten-Dienstobjekte repräsentierten Kategorien und/oder Dienste an den Nutzer eine erneute Spracheingabe des Nutzer empfängt, welche vom Sprach-Dialogsystem extrahierbare neue Suchworte enthält, auf Basis der neuen Suchworte in einer zweiten Suche eine neue Anzahl von

20 Kandidaten-Knotenpunkte und/oder Kandidaten-Dienstobjekte ermittelt, und die Schnittmenge der bei der ersten Suche ermittelten Kandidaten-Knotenpunkte und/oder Kandidaten-Dienstobjekte und der bei der zweiten Suche ermittelten Kandidaten-Knotenpunkte und/oder Kandidaten-Dienstobjekte bildet, und die Kategorien und/oder Dienste, welche durch die in der Schnittmenge befindlichen

25 Kandidaten-Knotenpunkte und/oder Kandidaten-Dienstobjekte repräsentiert werden, bevorzugt an den Nutzer ausgibt.

## 11. Automatisches Sprach-Dialogsystem (1)

- mit einer Spracherkennungseinrichtung (2) und einer Sprachausgabereinrichtung (3) zur Kommunikation mit einem Nutzer
  - mit mehreren vom Nutzer auswählbaren Diensten (9) innerhalb des Sprach-Dialogsystems (1) und/oder Mitteln (7) zur Vermittlung des Nutzers an vom Nutzer auswählbare Dienste (10) durch das Sprach-Dialogsystem (1),
  - mit einer Dialog-Kontrolleinrichtung (4) zur Steuerung des Dialogs zum Auswählen eines Dienstes (9, 10) durch den Nutzer
  - und mit einer Datenbank (6) mit einer hierarchischen Datenstruktur (DS) mit einer Mehrzahl von Knotenpunkten (K) und einer Mehrzahl von Pfaden (P) zur Verbindung der Knotenpunkte (K) untereinander und zur Verbindung von Knotenpunkten (K) mit Dienstobjekten (D), welche jeweils an einem Ende eines Pfades (P) in der Datenstruktur (DS) angeordnet sind, wobei die Dienstobjekte (D) die zur Verfügung stehenden Dienste (9, 10) repräsentieren und die Knotenpunkte (K) jeweils Kategorien repräsentieren, in welche andere Kategorien und/oder Dienste, die durch die innerhalb der hierarchischen Datenstruktur (DS) auf einer Ebene (II, III) unter dem jeweiligen Knotenpunkt (K) angeordneten weiteren Knotenpunkte (K) und/oder Dienstobjekte (D) repräsentiert werden, klassifiziert sind,
- 20 dadurch gekennzeichnet
- dass zumindest zu einem Teil der Dienstobjekte (D) und/oder Knotenpunkte (K) innerhalb der Datenstruktur (DS) eine Mehrzahl von verschiedenen Pfaden (P) führt, dass jedem Knotenpunkt (K) und jedem Dienstobjekt (D) der Datenbank (6) ein oder mehrere Schlüsselwörter (S) zugeordnet sind,
- 25 und dass das Sprach-Dialogsystem (1)
- eine Analyseeinheit (11), um bei einem Empfang einer Spracheingabe des Nutzer aus dieser Spracheingabe Suchwörter zu extrahieren,
  - eine Sucheinheit (12), um auf Basis der Suchwörter eine Anzahl von Kandidaten-



- 5 Knotenpunkten (K) und/oder Kandidaten-Dienstobjekten (D) innerhalb der Datenbank (6) zu suchen, deren zugeordnete Schlüsselworte (S) gemäß einem vorgegeben Akzeptanzkriterium mit den Suchworten übereinstimmen, wobei die Sucheinheit (12) derart aufgebaut ist, dass sie eine Suche in mehreren
- 10 Suchschritten durchführt, bis nach einem Suchschritt die Anzahl der ermittelten Kandidaten-Knotenpunkte (K) und/oder Kandidaten-Dienstobjekte (D) oberhalb einer vorgegebenen Minimalanzahl und unterhalb einer vorgegebenen Maximalanzahl liegt,
- 15 - und eine Promptgenerierungseinheit (13), um nach einem erfolgreichen Suchschritt ein Sprachausgabe-Menü zu generieren, um die durch die ermittelten Kandidaten-Knotenpunkte (K) und/oder Kandidaten-Dienstobjekte (D) repräsentierten Kategorien und/oder Dienste (9, 10) an den Nutzer zur Auswahl einer bestimmten Kategorie oder eines bestimmten Dienstes (9, 10) mittels der Sprachausgabeeinrichtung (3) anzusagen,
- 20 aufweist.

12. Computerprogramm mit Programmcode-Mitteln, um alle Schritte eines Verfahrens nach einem der Ansprüche 1 bis 10 auszuführen, wenn das Programm auf einem Computer eines Sprach-Dialogsystems ausgeführt wird.

PHDE020059

1/3

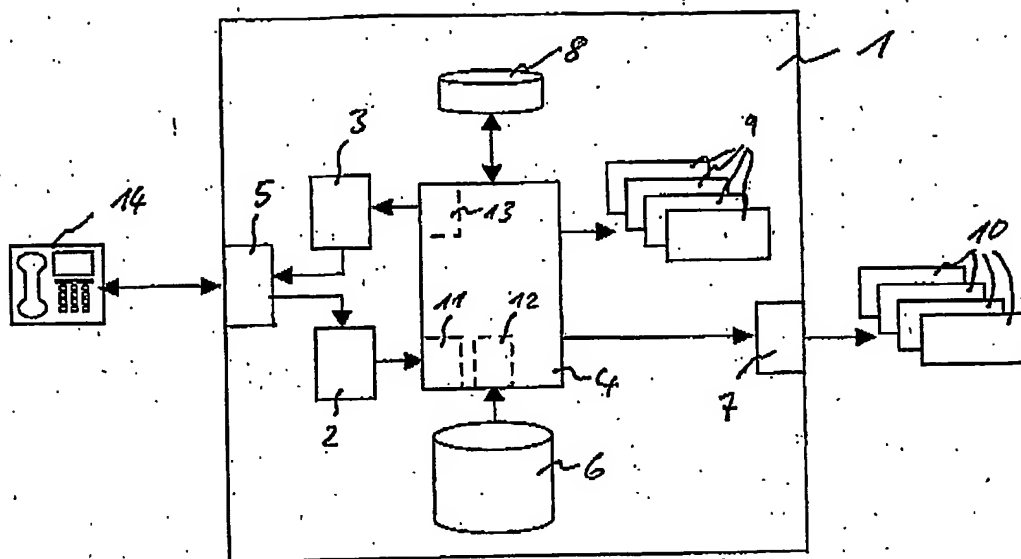


Fig. 1

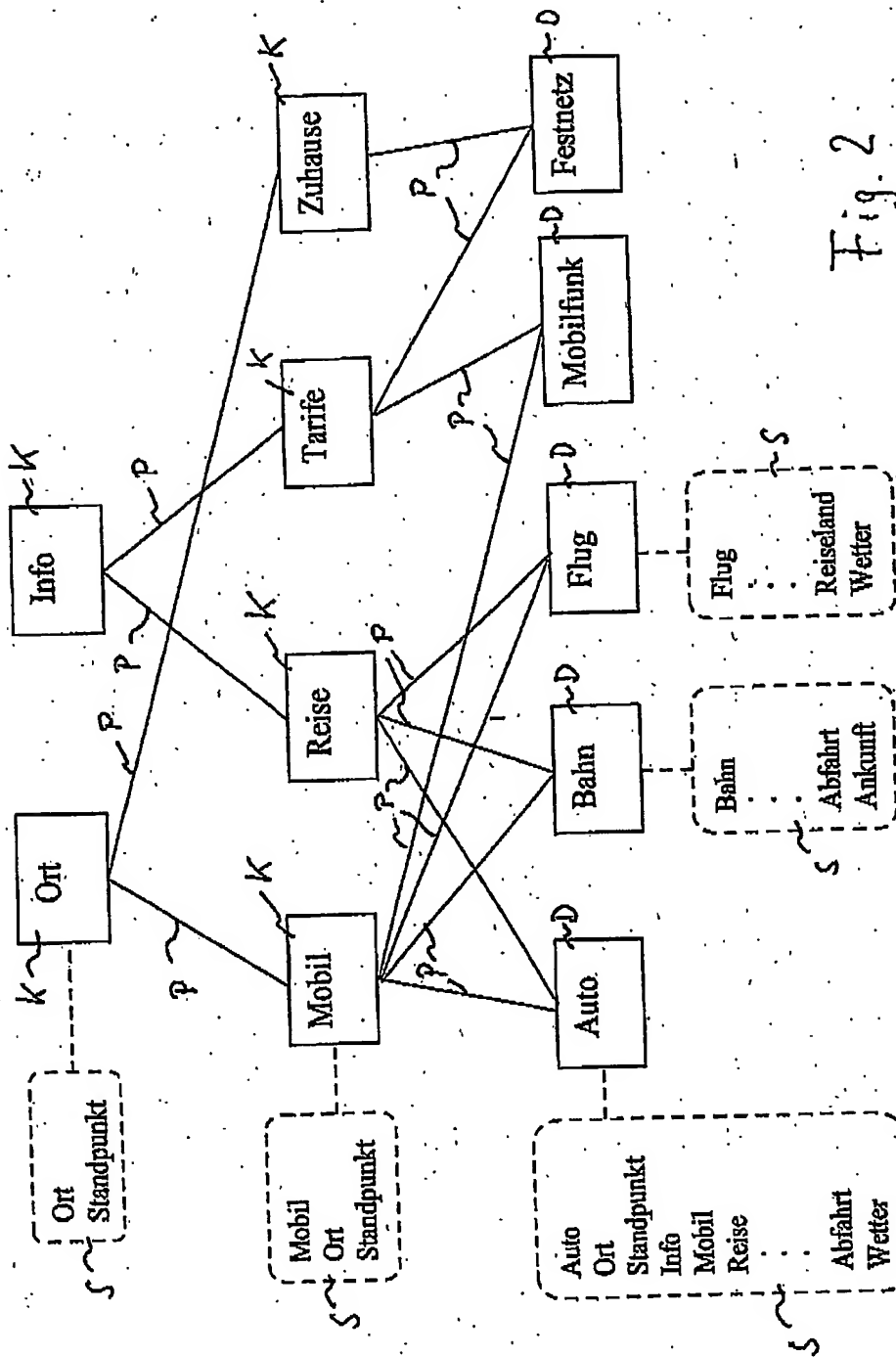


Fig. 2

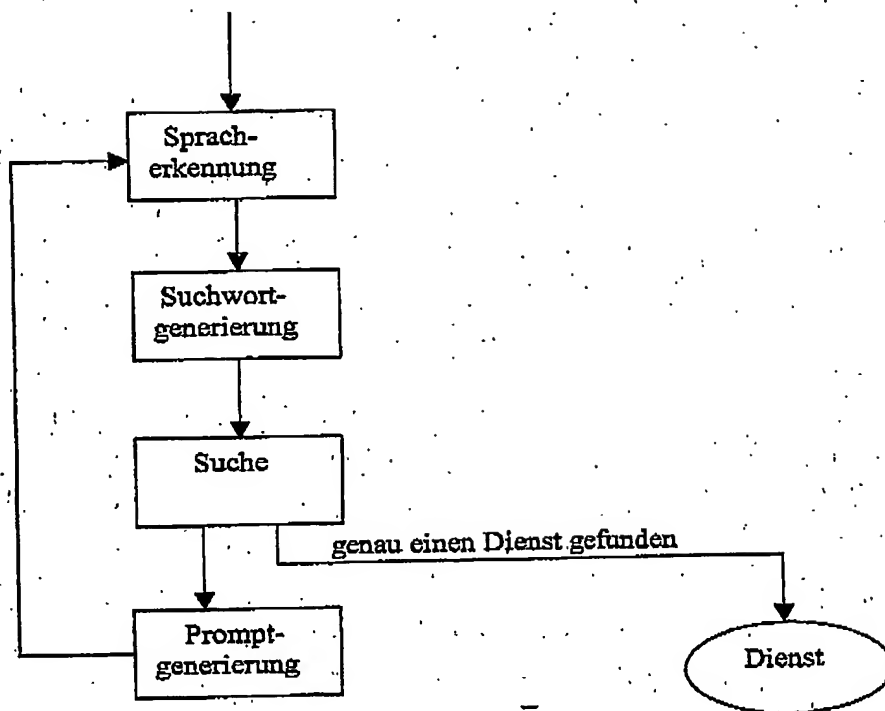


Fig. 3

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**